Some Considerations on Dependability Issues and Cyber-Security of Cyber-Physical Systems

Jean-Marc Thiriet
Univ. Grenoble Alpes, CNRS, Grenoble
INP, GIPSA-lab
38 000 Grenoble
jean-marc.thiriet@univ-grenoble-alpes.fr

Stéphane Mocanu
Univ. Grenoble Alpes, CNRS, Grenoble
INP, LIG
38 000 Grenoble
stephane.mocanu@grenoble-inp.fr

Abstract—For the last recent years, there has been a convergence between computer engineering approaches and automation aspects (industrial systems, internet of things) also called cyber-physical systems, for the development of process-based cyber-security strategies. Classically, security studies are based on risk analysis. Compared to classical IT approaches, the actual process (for instance a nuclear power plant or a chemical process) or system (autonomous car, drone) are taken into account in our approach for two reasons. The first reason is that the vulnerabilities of such systems or processes vary dynamically as a function of the time, the second reason is because the “standards” context is depending on the application domain and relationships with the IEC 61508 functional safety standard seems relevant. The paper presents a state of the art of problematics and proposed some approaches to these issues.

Keywords—cyber-security, industrial systems, internet of things, internet of things, functional safety

I. INTRODUCTION

This paper presents a state of the art of problematics dealing with cyber-security issues for cyber-physical systems. The main problematic of such systems is the impact. For classical IT (Information Technology) systems, the impact is in the digital field, it can be a denial of service, the stealing of data (violation of confidentiality), the destruction of systems or data (violation of integrity), but the impact is generally not in the physical world. For cyber-physical systems such as “large” industrial systems, smart grids, power plants or “embedded” “things” (embedded objects; drones, autonomous vehicles…) the impact of a malware or attack maybe to control remotely the object to divert it from its mission or to make it behave in a dangerous way: in both cases, catastrophic consequences may occur in the physical world (accident, death…).

During many years, industrial systems have not been cyber-protected as such, considering that the communication and processing architectures of such systems did not require to be taken into account. The main reasons were the heterogeneity of architectures (specific PLCs (Programmable Logic Controllers) and Operating systems, specific industrial networks) completely different of the ones used for Information Systems (classical computers and servers, TCP/IP networks). For this reason the main security strategy was based on “obscurancy” considering the fact not to be connected on a network as enough to be cyber-protected. The emergence of specific malwares such as Stuxnet discovered in 2008 [1] has changed the philosophy of cyber-security of industrial systems: Stuxnet was a Trojan-based virus exploiting four 0-day vulnerabilities (0-day means non detectable and with no possible protection mechanism) concerning propagation with a USB key for the first one, the print spooler of Windows for the second one and the last 2 vulnerabilities allowing increasing privileges to reach the administrator level to change code in the SCADA (Supervisory control and data acquisition) systems. The purpose of Stuxnet was to attack a precise configuration of the Simatic system used in some PLCs. The final target was a nuclear power plant in a tierce country. 30 000 computers were infected at the time… Flame is another malware discovered in May 2012 [2]. HavexRAT [3] was also a Trojan-based virus, it was an attack done against the energy sector in a tierce country. One of the purposes was to gain access to devices used for ICS (Industrial Control System)/SCADA. Many of the recent high-profile security incidents such as the Stuxnet worm [4] and the CrashOverride [5, 6] attacks are instances of such process-oriented attacks [7].

Embedded systems (internet of things) are also another important aspect nowadays: they invade everyday life: smartphones, connected toys, augmented reality. More complex systems are also present with more and more computers or digital electronics (automotive, household appliances), including critical functions as it is usually the case in transport [8]. The way to protect such systems, which can be vulnerable due to the difficulty of implementing locally protection programmes, and also because their impact may be severe (accident of an autonomous car due to hacking), is also at this time a real challenge.

In the remaining we define cyber-physical systems as systems which are controlled and/or monitored in the digital world; the controlled system is in the physical world, it can be a large process or an embedded system.

In the paper we refer to cyber-security as the possibility to circumvent the normal functioning of systems, by mean of actions in the digital world. It means that we don’t take into account physical attacks such as direct destruction or direct electromagnetic jamming for instance.

The next section presents the concepts of dependability and functional safety. Dependability is important from our point of view since cyber-security issues to cyber-physical systems may bring to a catastrophe. Functional safety was developed in order to take account of the evolutions of technology, in particular the use of embedded micro-controllers and communication networks. Section III deals with risk analysis. To envisage protection means to know from which danger or unexpected event we want to be protected against. Risk analysis allows the classification of risks according to their occurrence probability and severity, followed by an action plan defining what are the priorities in terms of security. Section IV explains what are the steps to respect in order to have an attack to succeed and give some inputs about cyber-physical systems. Section V proposes an
approach based on IDS to protect more generally communication and information infrastructures and more specifically industrial systems. Section VI gives some information about various network protocols which are used in the field of cyber-physical systems. In section VII are discussed some aspects relative to security audits with some specific view about cyber-security of industrial plants.

II. DEPENDABILITY AND FUNCTIONAL SAFETY

Dependability or RAMSS is generally considered as composed of four attributes: Availability, Reliability, Maintainability, Security and Safety) [9]. The main purpose of dependability is to be able to give an estimation, a probability for the system under study to behave as expected as far as these RAMSS attributes are concerned. This approach is used in many fields, particularly for many years it has been used in mechanical engineering, thermal engineering, analog electronics... The widespread of digital-based technologies (in the 90s micro-controllers and DSPs in order to allow the design of smart or intelligent devices or systems, communication networks) changed the paradigm of dependability because such systems are dynamical systems with a more complex architecture than the traditional series or parallel architectures. The development of the dependability issues for such systems was more difficult [10].

Functional safety, in particular around the IEC-61508 standards [11] and its derivative gave a frame to study this new category of digital-based systems. Functional safety of cyber-physical systems is a delicate subject, both scientifically and on the societal dimension [8]. Evaluating functional safety consists in giving a value to quantify the level of risk associated with the use of such an architecture according to the mission that it should achieve. This level of risk is characterized by the consequence on the human or material environment of dangerous situations, as well as by the probability or possibility of occurrences that such a situation may occur. At the societal level, automated or computerized systems are present everywhere and their relative risks should be studied.

Cyber-physical systems are composed of communication networks, embedded smart sub-systems, distributed diagnosis functionalities. If we consider the communication aspect, [10] traditionally two types of approaches are proposed: “classical” approaches in which the communication network is totally ignored [12] and network-centric approaches where the network performance is studied without taking account of the application [13] [14]. In “classical” approaches, faults are considered as permanent nature, and the system failure is defined as a function of the failure of the components and their logics of operation. In network approaches, the failure is defined as the non-respect of the delivery time of a message. The use of networks and digital electronic components in replacement of some mechanical parts (eg X-by-Wire in vehicles) or analog electronic devices makes the system more sensitive to disturbances and requires specific protections. Transient faults may occur in such systems and have an impact on the reliability of the system. This transient dimension is interesting from the point of view of cyber-security because they can be considered as transient vulnerabilities which as a consequence may be considered as threats, or contextual threats.

In previous works [10] we have shown there is a relationship between the reliability of the whole system and the probabilities of delays or loss of messages in the network. One difficulty with this type of study is that faults affecting the network do not always have the same effect; for example in a closed loop distributed around a network loss of messages in the transitional phase does not have the same effect as during the steady state. We are in a situation of “dynamic reliability” because the relationship between the system failure and failures of components change over time. In this previous works, we have proposed a modelling of the system behavior considering the presence of a constant delay in the loop (robustness study, stability limit ...). In industrial reality, the network disturbances are random, the elements connected to the network function asynchronously and several message exchanges may be necessary in the same sampling period. The consequence is that in general an analytical study is hardly feasible. The ability of control systems to compensate the effects of certain component failures (like the network) leads to a redefinition of the concept of system failures. The consequence is that the assessment of reliability of the system is dependent on the functional evaluation and becomes more difficult if not impossible with the traditional methods. To overcome these difficulties, an approach combining modelling (based on Petri nets, Stochastic Activity Networks) and simulation may be used and we proposed some in the past [15, 16, 17]. The Monte Carlo approach is generally used for the statistical evaluation of the parameters dependability through many simulations (or stories).

III. RISK ANALYSIS

Risk analysis involves the development of an assessment of risk combinations through the collection and integration of information relating to scenarios [8], it is based classically on occurrence frequencies of events and consequences of these events if they occur. It is a major component of risk management in a company. It is generally admitted that a hazard and risk analysis must be performed from the design stage for the process. This analysis relates to the description of dangerous events, to the description of the consequences of these events, the determination of the requirements for the reduction of this risk, the assignment of safety to protective layers, etc. This risk analysis is intended to be an instrument of prevention and protection and in the form of methods involving phases of identification, evaluation and hierarchy [18].

The risk reduction must be achieved in order for the risk to be tolerable despite a dangerous situation. The purpose of determining the tolerable risk of a dangerous event is to indicate what is reasonable in relation to the frequency of the dangerous event and its consequences. When a risk is described as unacceptable, prevention methods are used in this case.

Accident prevention refers to the elimination of hazards or the reduction of the occurrence of unexpected events by improving the security of control or by the establishment of means preventing the appearance or propagation of dangers (such as Safety Instrumented Systems ...). Protection comes after the failure of prevention means and it relates to the mitigation of the consequences of an accident by means limiting the damages (emergency systems, emergency
levels of protection are designed to reduce the frequency of the hazardous event and its consequences.

In the field of cyber-security, three concepts are taken into account for cyber-security [19]:

- Asset: Represented by monetary value, they are considered as anything of worth that can be damaged, compromised, or destroyed by an accidental or deliberate action; a asset’s worth is generally far more than the simple costs of replacement (image, legal issues…).

- Threat: potential event that, if realized, would cause an undesirable impact; two factors plays in the severity of a threat: degree of loss and likelihood of occurrence.

- Vulnerability: absence or weakness of cumulative controls protection in a particular asset, it is estimated as percentages based on the level of control weakness.

To conclude on the aspect of the severity-probability relation really widespread in the field of dependability, we can complete by considering that in the field of cyber-security it is not easy to have a good estimation of the probability because it is strongly related to some specific contexts (cyber-war, economic competition…). If we consider attacks, the question is not just the probability of being a victim of an attack. The question is that we want to protect the assets from an attack considered as sufficiently probable, for which we will be interested in the difficulty of attack, linked to the probability for the attack to succeed.

IV. CYBER-ATTACKS

A cyber-attack is considered as a malicious action designed to impair security. An attack is the realization of a threat, and it needs a vulnerability exploit. An important aspect is that an attack could occur (and succeed) only if there is a vulnerability. In the field of cyber-security we used to define two types of targets for cyber-attacks. Convenient target means that the attacker is searching for resources and is not strongly interested by the target as such. Chosen targets are the real targets, which means this is the purpose of the attacker to reach this target in order for instance to steal information (violation of confidentiality) or to modify some data (violation of integrity). In the case of cyber-physical systems, because their architecture are generally more specific, these systems are more sensitive as chosen targets (more rarely as convenient targets), and more specially to the violation of integrity, because some attackers wants to change the configuration files or because of some attempts to try to upload corrupted software or firmware.

In order to achieve an attack on a chosen target, a hacker uses a strategy in 6 steps:

- Recognition and collection of information: Domain names, DNS servers, blocks of assigned IP addresses, public IP addresses, www, ftp, e-mail…, types of machines and OS on which the services are carried out, mechanisms available for the control of the access to the network, Type of firewall and IDS (Intrusion Detection System)... Cartography of the network, using SNMP, type of access connections

- Scan of the services and ports

- Enumeration: extraction of information on the valid accounts and the resources, network resources and shared resources, users and groups (as a function of the Operating system), appliances, character strings sent in response by the equipment

- Obtaining an access: use of the functionalities of the O.S., use of the functionalities of software, benefiting from a bad configuration, “Opened” system, default configuration (administrator name and password!), functionalities activated by default, scripts available on the system and sometimes activated by default (Unix/Linux), Hijacking SQL queries when querying a database via web interface, automated Attack (ex: scan of port 80 of a whole C-class block of addresses in order to seek a fault)

- Extension of the acquired privileges : to carry out code to obtain privilege, to seek to decipher other passwords, to scan for non ciphered passwords, to seek possible inter-network relations, to identify badly configured files or shared resources permissions

- Cover the traces: to dissimulate to the administrator the fact that one penetrated the system, to eliminate the entries (inputs) in the event logs and the registers, to empty the file of history.

In the field of cyber-physical systems, it needs specific hackers with the knowledge or experience about specific software, firmware or Operating systems used in such applications. From this point of view it is more difficult because there are at this stage less experts, but generally these systems are also not so much protected.

The main categories of attacks are denial of service, identity spoofing, protocol bypass, sniffing, man-in-the-middle.

For cyber-physical systems, the main aspect concerns the potential impact on the physical component, which may be the result of a classical IT attack but also of a more specific attack like using vulnerabilities in an industrial software, taking account of the stability limit of a controlled system (aircraft, vehicles) achieving a mission, changing the value of some sensors in order to provide the control algorithm a wrong view of its environment and situation, change values in the mission, upload of an unexpected software or firmware, etc… We may see from this list that some issues are strongly related to the classical IT world whereas some others are specific to cyber-physical systems and their functioning modes and contexts.

V. CYBER-PROTECTION OF INFRASTRUCTURES AND INDUSTRIAL SYSTEMS USING IDS PROBES

In order to protect IT infrastructures, we may use Intrusion Detection Systems. Two categories of IDSs exist. Network-based IDSs (NIDSs) reside on a discrete network segment and monitor the traffic on that segment. They usually consist in a network appliance with a network interface card (NIC) that is intercepting and analyzing the network packets in real time. NIC are generally in promiscuous mode, this is a « furtive » mode in order not to use any IP address. Host-based IDSs (HIDS) use small programs that reside on a host computer (web server, mail server…) in order to monitor the operating system and detect
inappropriate activities. Signature-based IDSs use a signature or attributes that characterize an attack which are stored for reference (if there is a match, a response is initiated). Statistical anomaly-based or behavior-based IDSs dynamically detect deviations from the learned patterns of «normal» user behaviour and trigger an alarm when an intrusive activity occurs.

In paper [21] one of the problematic is what is the better way to locate the IDS (Intrusion Detection Systems) probes. If the processing is centralized, the probes are not considered exactly as IDS but a minimal processing is achieved at the level of the probe and a centralized IDS is in charge to achieve the analysis. It is possible this way to achieve a correlation between several points in the network. In this case the centralized IDS should be able to process the complete traffic of the industrial process. Another approach consists in using probes as complete IDSs analyzing locally the traffic and sending only alerts at the higher level. To this aim the authors have integrated Modbus within Suricata as a proof of concept. In order to generalize the work, it is necessary to achieve the same type of works on other protocols such as Profinet.

If we consider the use of IDS in the field of Industrial Control Systems, the following classification has been provided by [7]. In the cyber domain, the communication-based approaches [22, 23, 24, 25] are interested in the vocabulary and grammar of industrial protocols, these approaches are based on the knowledge we have about the protocol mechanisms and are really relevant in the case of deterministic processes. Flow-based approaches are closer to the approaches used in the IT field, [26, 22, 27, 28] focus on the detection of irregular flows within the ICS. Telemetry oriented approaches [28, 29, 30, 31] focus on building a base profile of network exchanges using statistical measures or classification models. Compared to communication-based approaches, intelligent nodes based approaches [32, 33, 34, 35] are relatively scarce. This can be explained by the specificity of the components in an ICS, the limited memory and computing resources, as well as instrumentation, for instance the authors in [32] develop an intrusion detection approach based on the analysis of the execution time of tasks in real-time devices. The other category is composed of the Physical domain oriented approaches which take into account incorrect behaviors at the level of the physical process without generating overtly abnormal behavior in the cyber domain, this type of IDS approaches incorporate more knowledge about the physical process [36, 37, 38, 39, 40, 41]. Finally, a number of works [42, 43, 44] cover multiple aspects of the above taxonomy. Such a wide coverage is motivated by the need to detect sophisticated attacks, identify accidental deviations, and reduce false positives.

VI. SOME CONSIDERATIONS ON THE NETWORK FOR CYBER-PHYSICAL SYSTEMS

Networks are an important part of cyber-physical systems and these networks may be really various as a function of the application field such as classical TCP/IP networks, industrial networks, wireless networks, mobile networks [45]. Another general division of the network could be as follows [46]:

- deterministic networks,
- non-deterministic networks.

In deterministic networks we can foresee the time of transmission of the required amount of data as well as delay appeared on the network. An example of deterministic network is WorldFIP. In case of non-deterministic network there are random delays and we cannot determine the exact time of transmission. The common communication network could be represented by Ethernet. Ethernet uses CSMA/CD (Carrier Sense Multiple Access / Collision Detection) [standard IEEE 802.3] [47] mechanism for access on communication medium. This mechanism is realized at the 2nd layer of ISO/OSI model [48]. The main principle of this kind of access mechanism is collision detection by node when the information is sent. When another signal is detected on the medium, then the transmission of the packet is interrupted and a jam signal is sent. At this moment, each node is waiting for a random time interval. The performance parameters are sufficient when the network loading is low. Increasing the amount of data which should be transmitted on the network (network loading) increases, in the same time, the number of collisions. Thus, the corrupted packets must be retransmitted and the network is gradually overloaded.

Ethernet network with CSMA/CD mechanism has worse performance parameters in comparison with industrial networks, when the network is high loaded. For example, with the same bit rate as in technological networks, the Ethernet network is not able to transmit required data within the required time intervals [49] [50]. Ethernet is a type of non-deterministic network. This feature is given by the inability to foresee delays caused by the network due to its medium access method. In order to prevent undesirable features and increase the dependability of the whole Network Controlled System (NCS), special networks were designed for industrial applications. There are several types of used control networks:

- CAN, CANopen, DeviceNet – represents the network with CSMA/CR [48] or CSMA/AMP [51] medium access method,
- PROFIBUS, ControlNet, FlexRay – represents token passing networks with TDMA medium access method,
- Industrial Ethernet and other modifications such as EtherCAT or Ethernet PowerLink (EPL).

The difference in approach is covered by different access mechanisms for the nodes connected to the network. Generally, the most known technological networks apply CSMA/AMP (CAN) or TDMA (token passing bus / ring) medium access method. There are some modifications based on conventional communication network structure. Popularity of the classical Ethernet pushes it in industrial application. In order to be able to accomplish requirements given by its application within NCS, some features have to be modified. The choice of a suitable type of control network (deterministic or non-deterministic) depends on network performance parameters given under different conditions. An important aspect is the dynamic of the controlled system as well as the complexity of the systems, number of components connected to the network and others which should not be ignored when the NCS is designed.

As far as cyber-security is concerned, we can consider that a deterministic network is generally more robust or resilient to denial of service attacks for instance, because we have a control of the temporal aspects. Temporal aspect is an
important issue for cyber-physical systems and so a potential source of vulnerabilities.

Until recently [52], few industrial protocols have proposed security mechanisms (integrity control more robust than a CRC, authentication or data encryption). Among protocols proposing security mechanisms, there are DNP3, OPC UA, IEC 61850. DNP3 is the secured version of DNP3, allows the authentication of devices and to ensure data integrity by using certificates, hash functions and TLS (Transport Layer Security). OPC UA (OLE for Process Control Unified Architecture) is based on being used by all the operational components of an ICS (PLCs, HMI, SCADA servers and clients, and any component using an OPC server or client). Contrary to other OPC protocols, OPC UA is not based on DCOM for various reasons, among which the security of the protocol. OPC UA contains mechanisms to ensure authentication, confidentiality and integrity of exchanged data. IEC 61850 is a standard specifically devoted to Smart Substation Automation Systems. Standards take account of cyber-security of exchanges, by using TLS, hash and certificates. We have shown in previous works [53] how IEC 61850 protocol is not secured and how it is possible to bypass in particular some temporal restrictions.

VII. Audit

Audits are used to analyse [19] the strength, weakness of organisations, plants, institutions... In this paper we are interested about cyber-security audits. This section reminds what is a cyber-security audit, and we will give some inputs for cyber-security of industrial plants.

Cyber-security IT audits are based in the IEC 27002 standard [54]. An audit is a specific procedure, this procedure should be agreed by the parties and documented clearly. The limits of the audit should be strongly defined. The audit is composed of several parts. The first part is dedicated to the security policy, the saving policy, the disaster recovery. Several parts are then dedicated to the various parts of the information system: data centers, servers, rooters, firewalls, operating systems, databases, applications, WLAN, mobile devices.

For industrial systems [52], ANSSI (Agence nationale de la sécurité des systèmes d'information, https://www.ssi.gouv.fr/) in France has published in 2014 some references for industrial systems. ANSSI has defined some requirements audit providers should respect to audit “OIV” (Opérateurs d’intérêts vitaux). Cyber-security audits for industrial plants should be conducted according to standards, either generic (ISO 27000) or specific (NIST 800-82, NERC CIP for production and transport of electricity, IEC 62443) but also using the ANSSI referential ANSSI14 which will be compulsory for “OIV”.

Finally, for IT as well as for Industrial architectures, pen tests may be achieved to test the conformity of the installation to the expected level of cyber-security.

VIII. Conclusions

The purpose of this paper is to give a state of the art and problematics of cyber-security of cyber-physical systems. Due to the cyber-physical aspects with impact in the physical world, the link with dependability and safety, especially functional safety, is natural even if the methodologies are not completely available yet. Risk analysis is the tool classically used in the field of dependability to have an estimation of the risks, their probabilities and severities. Cyber-threats are then presented, in order to understand how a system can be attacked to well protect it. Some general aspects on how to protect the infrastructure are then proposed. Properties of various network protocols are then discussed. The last section deals with security audit, which is an approach to be used to get an objective evaluation, from a tierce-party, about the actual level of security of a system.

Some other issues need to be taken into account, in particular the issue of embedded systems and the way to implement cyber-security resources in autonomous embedded systems with limited calculation and sometimes energy resources.

As a whole, this question of cyber-security of cyber-physical system remains an open problem. Methodologies such as the one of the 27005 standards [55] may be used as a framework for the cyber-physical systems, but the knowledge about the controlled process, its behaviour, the potential vulnerabilities of the behavior, are some points which should be taken into account, which means a convergence between IT specialists and automation experts.

ACKNOWLEDGMENT

The authors would like to thank the PhD students which worked with us these last years on these subjects: Rony Ghostine, Abdelhak Mkhida, Jan Galdun, Maëlle Kabir-Querrec, Insa Sassi, Oualid Koucham, Tran Trun Duc.

REFERENCES


